UNIVERSITY OF ILLINOIS BULLETIN
ISSUED WEEKLY
VOL. XXVI FEBRUARY 5, 1929 No. 23
[Entered as second-class matter December 11, 1912, at the post office at Urbana, Illinois, under the Act of August 24, 1912. Acceptance for mailing at the special rate of postage provided for in section 1103, Act of October 3, 1917, authorized July 31, 1918.]

BULLETIN NO. 44

BUREAU OF EDUCATIONAL RESEARCH
COLLEGE OF EDUCATION

HOW PUPILS SOLVE PROBLEMS IN ARITHMETIC

By
WALTER S. MONROE
Director, Bureau of Educational Research

PRICE 50 CENTS

PUBLISHED BY THE UNIVERSITY OF ILLINOIS, URBANA
1928

The Bureau of Educational Research was established by act of the Board of Trustees June 1, 1918. It is the purpose of the Bureau to conduct original investigations in the field of education, to summarize and bring to the attention of school people the results of research elsewhere, and to be of service to the schools of the state in other ways.

The results of original investigations carried on by the Bureau of Educational Research are published in the form of bulletins. A list of available publications is given on the back cover of this bulletin. At the present time five or six original investigations are reported each year. The accounts of research conducted elsewhere and other communications to the school men of the state are published in the form of educational research circulars. From ten to fifteen of these are issued each year.

The Bureau is a department of the College of Education. Its immediate direction is vested in a Director, who is also an instructor in the College of Education. Under his supervision research is carried on by other members of the Bureau staff and also by graduates who are working on theses. From this point of view the Bureau of Educational Research is a research laboratory for the College of Education.

BUREAU OF EDUCATIONAL RESEARCH
College of Education
University of Illinois, Urbana

BULLETIN NO. 44

BUREAU OF EDUCATIONAL RESEARCH
COLLEGE OF EDUCATION

HOW PUPILS SOLVE PROBLEMS IN ARITHMETIC

By

WALTER S. MONROE
Director, Bureau of Educational Research

PUBLISHED BY THE UNIVERSITY OF ILLINOIS, URBANA
1929

UNIVERSITY
OF ILLINOIS
PRESS

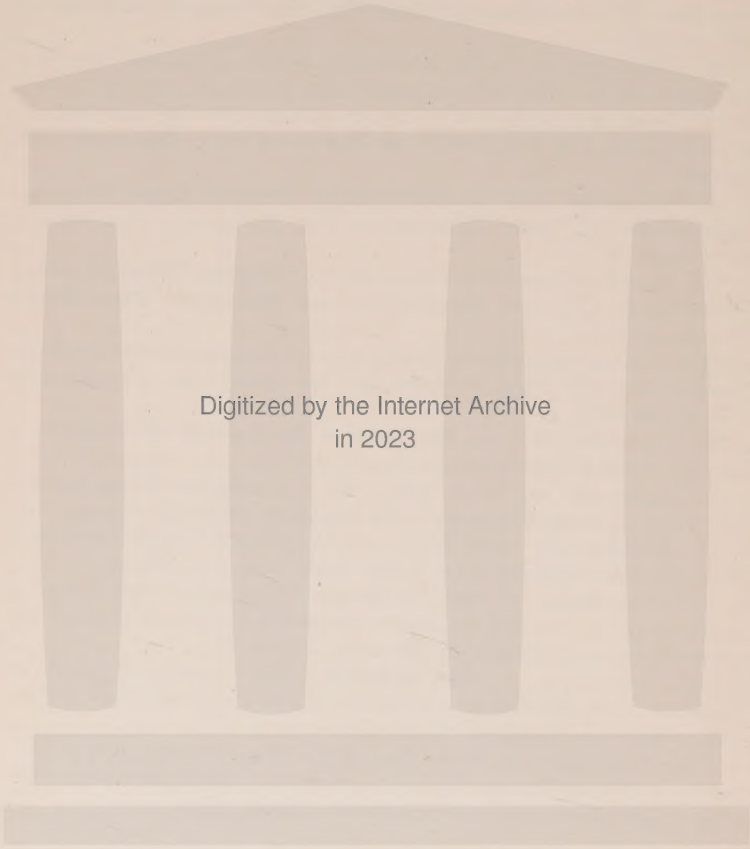
PREFACE

The problem studied in the investigation reported in this bulletin is an important one. Arithmetic is generally thought of as affording a large portion of the opportunities for reflective thinking in the elementary school, and it has been assumed that much of the training which pupils receive in this process is secured by solving arithmetical problems. If it is true, as the present investigation indicates, that pupils do not think reflectively in solving problems, it is obvious that their learning in the field of arithmetic is not what it is assumed to be. The reader of this bulletin, however, should bear in mind that the investigation deals only with the way in which pupils now solve problems. No attempt was made to ascertain if pupils could be taught to solve problems by thinking reflectively.

The data for the investigation were collected during the school year of 1926-27. Mr. John A. Clark, Assistant in the Bureau of Educational Research, supervised the scoring of the test papers. The tabulation of the data, however, was not completed until after Mr. Clark left the Bureau of Educational Research. Work on other projects during the school year of 1927-28 delayed the completion of the report.

WALTER S. MONROE, *Director*

LIST OF TABLES



HOW PUPILS SOLVE PROBLEMS IN ARITHMETIC

The problem of this study. In another place, the author has described the general process of solving verbal problems in arithmetic as follows:

> Meanings are connected with words and symbols; the implied question concerning a functional relationship is identified and answered; denominate number facts are recalled; numbers are read and copied. However, the response cannot be adequately described by enumerating the responses to such elements. Reflective thinking is involved. It should be noted that the total response to a verbal problem includes the determination of the calculations to be performed plus the response to the example formulated. Reflective thinking is involved in only the first phase of the total response.[1]

Elementary-school teachers and other observers of the performances of pupils in the field of arithmetic probably would criticise this statement by pointing out that it describes how pupils should solve problems, rather than how they do solve them. It is the purpose of this study to answer certain questions relating to the actual procedure that pupils follow in solving problems in arithmetic.

It seemed to the writer that the nature of pupils' responses to changes in the statement of a problem would furnish an indication of how they proceed in solving it.[2] Consequently, a series of tests was devised which made possible a comparison of responses of pupils to certain types of statements of the same problem. A statement of a problem was considered abstract if there was no reference to the activity in which it occurred. If this activity was indicated in the statement, it was called concrete. The terminology of a problem was considered technical when the terms used were those that are commonly employed by specialists in a particular field. The terminology was considered to be simple when it consisted of words and phrases in general use. A third modification in the statement of a problem was secured by introducing irrelevant data. The way in which these variations were introduced in the statement of certain problems is described later under the heading of the construction of the tests.

---

[1] Monroe, W. S. and Clark, J. A. "The Teacher's Responsibility for Devising Learning Exercises in Arithmetic," *University of Illinois Bulletin*, Vol. 23, No. 41, Bureau of Educational Research Bulletin No. 31. Urbana: University of Illinois, 1926, p. 22.

[2] The word "problem" is used throughout this bulletin to designate what is sometimes called a verbal problem. Arithmetical exercises frequently designated as examples are not included.

The specific questions for which information was sought may be stated as follows:

1. What is the difference between the responses of pupils to problems stated abstractly and the same problems stated in concrete terms?

2. What is the difference between the responses of pupils to problems stated in simple terms and the same problems stated in technical terminology?

3. What is the difference between responses of pupils to problems in which all the data given are needed for solving the problem and the same problems stated so as to include some irrelevant data?

After the data had been collected, a fourth question was added:

4. In solving a problem incorrectly in principle, what procedure do pupils follow?

In these questions, the phrase "responses of pupils to problems" refers only to the principle or plan of solution throughout the study. No attention was given to errors of calculation.

The construction of the tests. Four separate tests were constructed—Test A, Test B, Test C, and Test D. With one exception,[3] the test problems were taken from texts designed for the seventh grade. The wording of these problems was varied, and the various forms of statements arranged as shown in Section I of Table I. Problems 1 and 7 were presented in the same form on all four tests. In Test A, Problem 2 is stated in simple terminology (S), all of the data given are relevant (R), and the setting is concrete (C). In Test B, technical terminology (T) is used, all the data given are relevant (R), and the setting is concrete (C). The difference in the statement of the problem is the change from simple terminology to technical terminology. In Test C, the problem is stated in simple terminology (S), although the data given are relevant (R), and the setting is abstract (A). In Test D, technical terminology (T) is used, irrelevant data (I) are included, and the setting is abstract (A). The four forms of Problem 2 are as follows:

Problem 2

Test A—SRC. During a sale, Smith and Company reduced the price of furnaces 20%. When the purchaser paid cash, they gave an additional 5% off the sale price. Mr. Jones bought a furnace at the sale and paid $551 cash for it. What price did Smith and Company originally ask for the furnace?

Test B—TRC. Mr. Jones was allowed successive discounts of 20% and 5% on the list price of a new furnace which he bought from Smith and Company. If Mr. Jones paid $551, what was the list price of the furnace?

---

[3]Problem 3 was not taken directly or indirectly from any text. However, it represents a very simple type of problem that has a high frequency of occurrence in all texts.

Test C—SRA. An amount was reduced 20%. After an additional reduction of 5% of the remainder was made, the final remainder was $551. What was the original amount?

Test D—TIA. A man borrowed $551 for 60 days at 7% to pay a bill on which successive discounts of 20% and 5% were allowed. What was the original amount of the bill?

The variations of the statements of the other problems may be ascertained by referring to Appendix A. The arrangement of the various forms of statement in the four tests is indicated in the first section of Table I.

Plan of securing equivalent groups. It is obvious that it would not be satisfactory to have the same pupils respond to two or more statements of the same problem. Hence, it was necessary to secure four groups of pupils that were equivalent in ability to solve the problems selected. The method of random sampling was employed as a means of securing these equivalent groups. In order to have each of the tests given to a random sample of pupils, the four tests were arranged in alternate order so that when distributed to the pupils in the class, the first, fifth, ninth, thirteenth, and so forth, would receive Test A; the second, sixth, tenth, fourteenth, and so forth, would receive Test B; the third, seventh, eleventh, fifteenth, and so forth, would receive Test C; the fourth, eighth, twelfth, sixteenth, and so forth, would receive Test D. Since the tests were to be given in a large number of classes, it seemed that this plan of sampling would provide equivalent groups.[3a]

The equivalence of the groups on which the comparisons were based was still further insured by means of a carefully planned arrangement of the different forms of statements in the several tests. The six types of problem statements considered are represented by the letters S, T; R, I; C, and A.[4] Each statement of a problem necessarily possesses three of these characteristics, one of each pair. In this investigation, each statement was formulated so that its characteristics were present in one of the following eight combinations: SRC, TRC, SIC, SRA, TRA, TIC, SIA, TIA. These various forms of statement were distributed among the different tests for each problem as indicated in Section I of Table I. This distribution was planned so as to provide the maximum data for answering the questions given on the preceding page. A study of Table I will reveal that no two comparisons of the same factors were made between the same groups of pupils.

[3a] These groups were equivalent not only in arithmetical ability but also in respect to teachers, textbooks, and other factors.

[4] *See* p. 8 and 20.

TABLE I. PROVISIONS FOR COMPARISONS OF PUPIL RESPONSES TO VARIOUS TYPES OF PROBLEM STATEMENTS[a]

I					II				III				IV			
Problem	Test				S compared with T				R compared with I				C compared with A			
	A	B	C	D	SRC TRC	SIC TIC	SIA TIA	SRA TRA	SRC SIC	TRA TIA	SRA SIA	TRC TIC	SRC SRA	TIC TIA	SIC SIA	TRC TRA
1	same	on all														
2	SRC	TRC	SRA	TIA	A–B								A–C			
3	TRA	SIA	SRA	TIC				A-C			B–C					
4	TIA	TIC	TRA	SIC		B–D				A–C				A–B		
5	SIA	SIC	SRC	TRC	C–D				B–C						A–B	
6	SIA	TRC	TIA	TIC			A–C					B–D		C–D		
7	same	on all														
8	SIC	SRA	TIC	SRC		A–C			A–D				B–D			
9	TRC	TRA	SIC	SRA				B–D								A–B
10	SRC	TRA	SIA	TIA			C–D			B–D						
11	TIC	SRC	TRC	TRA	B–C							A–C				C–D
12	SRA	TIA	SIC	SIA			B–D				A–D				C–D	

[a]Section I should be read thus: In Problem 2, Test A, are the factors SRC; in Test B, TRC, etc. In Problem 6, Test A, are the factors SIA; in Test B, TRC, etc.

Sections II, III, and IV should be read thus: S can be compared with T in problems where the other factors are kept constant, as in Problem 2, Tests A and B, where the factors RC are constant. In Problem 11, Tests A and C, the factors TC are constant; hence, factors R and I can be compared.

Provisions for making comparisons. Table I presents in summary fashion the provisions for comparing pupil answers to the various problem statements. Section II shows that provision for the comparison of the responses to simple terminology (S) with those made to technical terminology (T) was made in the case of each of ten problems.[5] In the case of Problems 2, 5, and 11, all of the data were relevant (R) and the setting concrete (C). In the case of Problem 2, the comparison is between Groups A and B; for Problem 5, it is between Groups C and D; and for Problem 11, it is between Groups B and C. In the case of Problems 4 and 8, the statement contained irrelevant data (I) and provided a concrete setting (C); the comparison in the former being made between Groups B and D, and in the latter between Groups A and C. The rest of the table should be interpreted in the same manner.

Pupils tested. The problems for the tests were taken from those that seventh-grade pupils are asked to solve, but it was thought wise to have the tests given also in a few sixth and eighth-grade classes. Usable data were secured from the different groups as follows: Sixth grade, 775; seventh grade, 5902; eighth grade, 2579; total 9256. These pupils represented forty-one cities in Illinois.[6]

The validity of data. The principal questions relative to the validity of the data are: (1) Are the problems representative of those that seventh-grade pupils are asked to solve? (2) Were equivalent groups secured? The judgment of the author determined the problems selected. As stated on page 8, they were taken from texts designed for the seventh grade. No analysis of these texts was made to determine the frequency of occurrence of various types of problems, but a previous study[7] had made the author familiar with the various types of problems in arithmetic texts. The selection was limited somewhat by the requirement that the problem be one whose statement could be varied according to the plans of the experiment. Although it was evident from the test papers that the pupils in several of the classes tested had not studied certain of the types of problems represented at the time the tests were given, the author believes that the problems are sufficiently representative of those assigned in the seventh grade for the purpose of this experiment.

---

[5]As has been stated previously, Problems 1 and 7 were presented in the same form on all tests.

[6]These cities are: Beardstown, Belvidere, Benton, Blue Island, Canton, Centralia, Champaign, Chicago, Chicago Heights, Cicero, Clinton, Danville, DuQuoin, Edwardsville, Elmhurst, Evanston, Forest Park, Granite City, Harvey, Herrin, Jacksonville, Joliet, Kewanee, LaGrange, Lyons, Madison, Marion, Moline, Mt. Carmel, Oak Park, Oglesby, Ottawa, Pekin, Quincy, Rockford, Rock Island, Streator, Taylorville, West Frankfort, Westville, Zeigler.

[7]Monroe, W. S. and Clark, J. A. "The Teacher's Responsibility for Devising Learning Exercises in Arithmetic," *University of Illinois Bulletin*, Vol. 23, No. 41, Bureau of Educational Research Bulletin No. 31. Urbana: University of Illinois, 1926. 92 p.

The second question, regarding the equivalence of the groups tested, was discussed on page 9.

Simple versus technical terminology. The data relative to the differences in responses, due to changing the terminology from simple to technical, are summarized in Table II.[8] The data for the different problems are not consistent. A few problems were attempted by slightly more pupils and larger per cents of the solutions were correct in principle when they were stated in technical terminology, but the opposite condition prevails in the majority of the problems. Considering the differences in the per cent of problems attempted and the differences in the per cent of problems correct in principle, only seven of the twenty are negative. Approximately the same conditions prevail in Tables VI, VII, and VIII, only twenty-one of the sixty differences being negative. Hence, in general, the data support the thesis that the use of what is called technical terminology increases the difficulty of arithmetical problems, but the exceptions indicate that such terminology does not necessarily make a problem more difficult. The most marked exception is Problem 3. Examination of the two statements of this problem shows that the essential difference between the simple and the technical statement is the use of "amount of a bill" in the latter in the place of "total cost" in the simple statement. Apparently, "amount of bill" was more easily understood by these pupils than "total cost." This probably indicates that "amount of bill" was more frequently used in the problems they had solved and hence a technical terminology becomes "simple" when it is used frequently. In other words, pupils as now taught learn to respond to certain words and phrases and when unfamiliar terms are substituted a larger proportion of the pupils make incorrect responses.

The effect of introducing irrelevant data. The facts relating to the effect of introducing irrelevant data in the statement of a problem are summarized in Table III.[9] Except in the case of Problem 4, this table indicates that the presence of irrelevant data makes the problem more difficult, although several of the differences are small. The largest difference in per cent of solutions correct in principle is for Problem 6. In this problem, the lengths of the three sides of a triangle were given and the irrelevant fact was the altitude. Apparently, many of the pupils, especially those in Grades VI and VII, had become so accustomed

[8]The data are given separately for the three grades in Tables VI, VII, and VIII in Appendix B.
[9]The facts are given separately for the three grades in Tables IX, X, and XI in Appendix B.

Table II. Variations in Pupil Responses to Problems When Technical Terminology (T) was Substituted for Simplified Terminology (S)—Grades VI, VII, and VIII

Problems Compared			Number of Pupils		Problems Attempted			Per Cent Attempted			Number Correct in Principle			Per Cent Correct in Principle		
Simple	Technical	Constant Factors	S	T	S	T	Diff.	S	T	Diff.	S	T	Diff.	S	T	Diff.
A 2	B 2	RC	2328	2311	1923	1937	− 14	82.6	83.8	−1.2	40	42	− 2	2.1	2.2	−0.1
C 5	D 5	RC	2313	2305	2093	1978	115	90.5	85.8	4.7	189	205	− 16	9.0	10.4	−1.4
B11	C11	RC	2311	2313	1135	1195	− 60	49.1	51.7	−2.6	27	66	− 39	2.4	5.5	−3.1
D 4	B 4	CI	2305	2311	1891	1526	365	82.0	66.0	16.0	1114	786	328	58.9	51.5	7.4
A 8	C 8	CI	2328	2313	1350	1463	−113	58.0	63.3	−5.3	431	333	98	31.9	22.8	9.1
A 6	C 6	IA	2328	2313	1846	1612	234	79.3	69.7	9.6	629	301	328	34.1	18.7	15.4
C10	D10	IA	2313	2305	1073	1052	21	46.4	45.6	0.8	95	74	21	8.9	7.0	1.9
D12	B12	IA	2305	2311	817	813	4	35.4	35.2	0.2	136	73	63	16.6	9.0	7.6
C 3	A 3	RA	2313	2328	2277	2284	− 7	98.4	98.1	0.3	784	1021	−237	34.4	44.7	−10.3
D 9	B 9	RA	2305	2311	1680	1669	11	72.9	72.2	0.7	871	745	126	51.8	44.6	7.2

TABLE III. VARIATIONS IN PUPIL RESPONSES TO PROBLEMS WHEN IRRELEVANT DATA (I) WERE SUBSTITUTED FOR RELEVANT DATA (R)—GRADES VI, VII, AND VIII

Problems Compared			Number of Pupils		Problems Attempted			Per Cent Attempted			Number Correct in Principle			Per Cent Correct in Principle		
Relevant	Irrelevant	Constant Factors	R	I	R	I	Diff.	R	I	Diff.	R	I	Diff.	R	I	Diff.
C 5	B 5	SC	2313	2311	2093	2019	84	90.5	87.4	3.1	189	122	67	9.0	6.0	3.0
D 8	A 8	SC	2305	2328	1556	1350	206	67.5	58.0	9.5	501	431	70	32.2	31.9	0.3
C 4	A 4	TA	2313	2328	1457	1449	8	63.0	62.2	0.8	577	740	−163	39.6	51.1	−11.5
B10	D10	TA	2311	2305	1175	1052	123	50.8	45.6	5.2	90	74	16	7.7	7.0	0.7
C 3	B 3	SA	2313	2311	2277	2099	178	98.4	90.8	7.6	784	683	101	34.4	32.5	1.9
A12	D12	SA	2328	2305	880	817	63	37.8	35.4	2.4	193	136	57	21.9	16.6	5.3
B 6	D 6	CT	2311	2305	1722	1585	137	74.7	68.8	5.9	1062	170	892	61.7	10.7	51.0
C11	A11	CT	2313	2328	1195	875	320	51.7	37.6	14.1	66	28	38	5.5	3.2	2.3

to making an addition response to a request to find the perimeter that they added the quantities given without noting what they were.

The introduction of irrelevant data in Problem 4[10] resulted in a larger per cent of solutions correct in principle. Tables IX, X, and XI reveal additional cases of this phenomenon. In Problem 4 the irrelevant data are furnished by enumerating the articles purchased. The presence of these items, together with the other changes in the phraseology, appears to have made the problem easier. The explanation of this phenomenon is, probably, to be found in the use of "total amount of a bill was $1500," in the place of "merchandise worth $1500." On page 12 it was pointed out that "amount of a bill" appeared to be a very familiar expression. If this is true, it is not difficult to understand why the pupils responded more satisfactorily when irrelevant data were included in the problem. Hence, the problem became easier, not because the irrelevant data were introduced, but because another change was made which more than overbalanced this change.

Abstract versus concrete terminology. Table IV[11] shows that, in general, a slightly larger per cent of pupils attempted the problems when they were given a real or concrete rather than an abstract setting, Problems 2 and 8 showing the greatest differences. The per cent of solutions correct in principle is approximately the same for the two types of statements, except in the cases of Problems 6 and 9. The differences for these problems are −8.0 and 8.2, respectively. Examination of these two problems does not suggest a probable explanation of this variation. It appears that the introduction of phrases descriptive of a concrete setting for a problem makes little or no difference in the responses of pupils. They respond to abstract problems about as satisfactorily as they do to those having a concrete setting.

Easiest form of problem the one pupils have learned to solve. The data of Tables II, III, and IV, and more especially those of the detailed tables in Appendix B, suggest to the writer that the terminology with which pupils have become familiar in their study of arithmetic is the easiest for them. Although this hypothesis could not be proved without information regarding the problem statements in the texts used by the pupils tested, it appears consistent with what we know about learning. If it is valid, this hypothesis implies that, in general, pupils do very little or no reflective thinking in solving arithmetic problems. Instead, they learn to make rather

[10] *See* p. 21 and 22 for the four statements of this problem.
[11] The data are given separately for the three grades in Tables XII, XIII, and XIV in Appendix B.

TABLE IV. VARIATIONS IN PUPIL RESPONSES WHEN LANGUAGE WHICH IS ABSTRACT (A) IN NATURE WAS SUBSTITUTED FOR LANGUAGE WHICH PRESENTS A REAL OR CONCRETE (C) SETTING—GRADES VI, VII, AND VIII

Problems Compared			Number of Pupils		Problems Attempted			Per Cent Attempted			Number Correct in Principle			Per Cent Correct in Principle		
Concrete	Abstract	Constant Factors	C	A	C	A	Diff.	C	A	Diff.	C	A	Diff.	C	A	Diff.
A 2	C 2	SR	2328	2313	1923	1747	176	82.6	75.5	7.1	40	43	− 3	2.1	2.5	−0.4
D 8	B 8	SR	2305	2311	1556	1387	169	67.5	60.0	7.5	501	441	60	32.2	31.8	0.4
B 4	A 4	TI	2311	2328	1526	1449	77	66.0	62.2	3.8	786	740	46	51.5	51.1	0.4
D 6	C 6	TI	2305	2313	1585	1612	−27	68.8	69.7	−0.9	170	301	−131	10.7	18.7	−8.0
B 5	A 5	SI	2311	2328	2019	2084	−65	87.4	89.5	−2.1	122	109	13	6.0	5.2	0.8
C12	D12	SI	2313	2305	858	817	41	37.1	35.4	1.7	142	136	6	16.6	16.6	0.0
A 9	B 9	RT	2328	2311	1659	1669	−10	71.3	72.2	−0.9	876	745	131	52.8	44.6	8.2
C11	D11	RT	2313	2305	1195	1174	21	51.7	50.9	0.8	66	67	− 1	5.5	5.7	−0.2

fixed responses to certain types of statement. Hence, when a different form of statement is used, they make no response or an inaccurate one. There are exceptions, of course; some pupils do reason in such situations. But the data collected in this investigation suggest that a considerable number do not think reflectively when they respond to arithmetic problems. This conclusion is reinforced by the data described in the following paragraphs.

Procedure in incorrect solutions. For the purpose of studying the incorrect solutions, 250 seventh-grade papers of each test were selected at random. The errors in principle were tabulated separately for each test. Table V gives the number of erroneous solutions, the number of different kinds of erroneous solutions, and the number of unique solutions; i.e., those which occurred only once. Considering the nature of the problems, the number of solutions incorrect in principle is surprisingly large, but the number of different kinds of erroneous solutions is probably more significant. Examination of the test papers revealed many solutions that seemed to have little or no rational basis. An attempt was made to identify the solutions in which there was evidence of a partial comprehension of the problem. Only 28 of the 253 erroneous solutions of Problem 2 were so classified. The remaining 225, or 89 per cent, were such that there seemed to be no evidence that the problem was even partially understood.

It may be argued that Problem 2, which involved successive discounts, was so difficult that it did not reveal the nature of pupil responses to problems. Problem 3, which seems to be relatively simple, was solved incorrectly in 80 different ways, but in only 18 of these was there evidence of partial comprehension. In view of the simple nature of this problem, it is astonishing that seventh-grade pupils should give 62 different solutions that contained no suggestion of even a partial comprehension of the problem. There were 115 different erroneous solutions for Problem 4, which was the easiest in the list, but in only 18 of these was there evidence of partial comprehension of the problem. It therefore appears that a very large proportion of the pupils tested made a response to the problem without understanding it. Since the comprehension of a problem is the first step in a rational solution of it, we have here additional evidence in support of the hypothesis that, in general, pupils do little or no reflective thinking in solving arithmetic problems.

General conclusion. Although the data of this investigation are not entirely consistent, they appear to substantiate the conclusion

TABLE V. NUMBER AND FREQUENCY OF ERRORS IN ONE THOUSAND TEST PAPERS

Problem	Erroneous Solutions					Kinds of Errors					Unique Errors			
	A	B	C	D	Total	A	B	C	D	Total	A	B	C	D
2	220	209	190	181	800	74	67	63	124	253	37	38	32	139
3	123	143	145	47	458	6	50	9	30	80	4	44	4	20
4	55	66	67	66	254	29	48	47	30	115	15	34	30	19
5	212	210	192	195	809	131	135	65	78	306	94	80	33	43
6	106	46	123	127	402	13	16	19	11	32	2	8	17	1
8	90	95	102	120	407	42	30	38	45	113	27	22	20	52
9	64	79	77	56	276	13	16	38	14	54	4	6	32	6
10	133	135	103	99	470	44	53	44	41	116	24	26	19	16
11	91	135	118	124	468	42	45	27	32	102	47	29	11	16
12	61	74	61	63	259	25	32	44	35	89	11	18	23	22

that a large per cent of seventh-grade pupils do not reason in attempting to solve arithmetic problems. Relatively few pupils follow the plan described on page 7. Instead, many of them appear to perform almost random calculations upon the numbers given. When they do solve a problem correctly, the response seems to be determined largely by habit. If the problem is stated in the terminology with which they are familiar and if there are no irrelevant data, their response is likely to be correct. On the other hand, if the problem is expressed in unfamiliar terminology, or if it is a "new" one, relatively few pupils appear to attempt to reason. They either do not attempt to solve it or else give an incorrect solution.

APPENDIX A

PROBLEMS USED IN TESTS

In the following pages are reproduced the various forms of the twelve arithmetic problems presented to more than nine thousand pupils of the sixth, seventh, and eighth grades in Illinois for the purpose of obtaining data in regard to the actual procedure that pupils follow in solving problems in arithmetic. Four different statements are given for each problem, with the exception of Problems 1 and 7, which appeared in the same form on all four tests. The characteristics of the problem statements are represented by symbols as follows:

S Simple terminology
T Technical terminology
A Abstract
C Concrete
R All data relevant to problem
I Irrelevant data included

Problem 1

Test A. Mr. Duncan drove his car 6375 miles one year and found at the end of the year that 340 gallons of gasoline had been consumed. At 21c a gallon, what was the cost of gasoline for one mile?

This same statement was used for Tests B, C, and D.

Problem 2

Test A—SRC. During a sale, Smith and Company reduced the price of furnaces 20%. When the purchaser paid cash, they gave an additional 5% off the sale price. Mr. Jones bought a furnace at the sale and paid $551 cash for it. What price did Smith and Company originally ask for the furnace?

Test B—TRC. Mr. Jones was allowed successive discounts of 20% and 5% on the list price of a new furnace which he bought from Smith and Company. If Mr. Jones paid $551, what was the list price of the furnace?

Test C—SRA. An amount was reduced 20%. After an additional reduction of 5% of the remainder was made, the final remainder was $551. What was the original amount?

Test D—TIA. A man borrowed $551 for 60 days at 7% to pay a bill on which successive discounts of 20% and 5% were allowed. What was the original amount of the bill?

In Test A, the technical term, successive discounts, was restated in simplified language. A real setting was added by using words such as sale, Smith and Company, and Mr. Jones. Only relevant data were used. Thus Problem 2, Test A satisfied the theoretical requirement of having simplified (S) terminology, relevant (R) data only, stated concretely.

In Test B, the technical terms, successive discounts, and list price, were retained. Relevant data only were used. The words, Mr. Jones, Smith and Company, and sale gave a real setting. Thus the combination of factors represented by the symbol TRC was satisfied.

In Test C, simplified terminology was substituted for the technical terms, successive discounts. Relevant data only were used. No attempt was made to give a real setting to the problem, but rather, the problem was stated abstractly. In the latter respect, Test C differs from Test A.

In Test D, the technical terms, successive discounts, were used. Irrelevant data, 60 days at 7%, were introduced and the problem did not present a real setting.

This procedure describes how the corresponding problems of the different tests were formulated.

Without describing in detail the variations in the remaining corresponding problems, each of the statements of corresponding problems will be listed under the problem, test, and combination the particular statement represents.

Problem 3

Test A—TRA. Find the amount of a bill for 3½ yds. of material at 40c a yd., and for 10 yd. at 85c a yd.

Test B—SIA. Find the total cost of 3½ yards of material, 3 inches wide, at 40c a yard, and 10 yards of material, 40 inches wide, at 85c a yard.

Test C—SRA. Find the total cost of 3½ yards of material at 40c a yard, and for 10 yards at 85c a yard.

Test D—TIC. Mrs. Henderson made the following purchases: 3½ yds. of ribbon, 3 inches wide at 40c per yd.; 10 yds. of curtain material, 40 inches wide, at 85c per yd. Find the amount of the bill.

Problem 4

Test A—TIA. The following articles were purchased: 5 vacuum cleaners, 12 parlor lamps, 7 rugs, and a suite of furniture. The total amount of the bill was $1500. Terms: 90 days, net; 60 days, 2%; 30 days, 5%; cash, 10%. Find the net amount if cash is paid.

Test B—TIC. The Hall Company buys 5 vacuum cleaners, 12 parlor lamps, 7 rugs, and a suite of furniture from Brown and Company, wholesalers, for a total of $1500. Terms: 90 days, net; 60 days, 2%; 30 days, 5%; cash, 10%. Find the net amount paid to Brown and Company if the Hall Company pays cash.

Test C—TRA. A dealer buys merchandise worth $1500. Terms: 90 days, net; 60 days, 2%; 30 days, 5%; cash, 10%. Find the net amount if cash is paid.

Test D—SIC. The Hall Company buys 5 vacuum cleaners, 12 parlor lamps, 7 rugs, and a suite of furniture from Brown and Company, wholesalers, for a total of $1500. The bill must be paid not later than 90 days from the day of purchase. 2% will be taken off if paid in 60 days; 5% off if paid in 30 days; or 10% off if cash is paid. Find the amount Brown and Company receives if the Hall Company pays cash.

Problem 5

Test A—SIA. A man purchased 50 articles at $1.50 each and sold them at $2.25 each. Allowing 20% of the selling price for expenses, how much was made on each article? This is what per cent of the selling price?

Test B—SIC. The Student Supply Store in a university district bought 50 fountain pens at $1.50 each and sold them at $2.25 each. Allowing 20% of the selling price for the running expenses of the store, how much was made on one fountain pen? This is what per cent of the selling price?

Test C—SRC. The Student Supply Store bought fountain pens at $1.50 each and sold them at $2.25 each. Allowing 20% of the selling price for the running expenses of the store, how much was made on one fountain pen? This is what per cent of the selling price?

Test D—TRC. The Student Supply Store purchased fountain pens at $1.50 each and sold them at $2.25. Allowing 20% of the selling price for the running expenses of the store, the profit is how much per fountain pen? The profit is what per cent of the selling price?

Problem 6

Test A—SIA. One side of a triangle is $3\frac{1}{2}$ ft., another side is $8\frac{1}{4}$ ft., a third side is $8\frac{3}{8}$ ft., and the distance from the top point to the base is $8\frac{1}{8}$ ft. What is the sum of the sides?

Test B—TRC. A pennant was cut so that its base was $3\frac{1}{2}$ ft., the top side was $8\frac{1}{4}$ ft., and the lower side was $8\frac{3}{8}$ ft. What was the perimeter of the pennant?

Test C—TIA. The base of a triangle is $3\frac{1}{2}$ ft. One side is $8\frac{1}{4}$ ft., the other $8\frac{3}{8}$ ft., and the altitude, $8\frac{1}{8}$ ft. What is the perimeter of the triangle?

Test D—TIC. A pennant was cut so that its base was $3\frac{1}{2}$ ft., the top side was $8\frac{1}{4}$ ft., the lower side was $8\frac{3}{8}$ ft., and its altitude $8\frac{1}{8}$ ft. What is the perimeter of the pennant?

Problem 7.

Test A. Walter sold his bicycle to Henry for $36 and gained 25% of the selling price. If Walter had wished to make 40% of the cost, what would have been the selling price?

This same statement was used in Tests B, C, and D.

Problem 8

Test A—SIC. A crew of 4 men working for 8 hours with a steam shovel, dug a basement 40 feet long, 24 feet wide, and 8 feet deep. Mr. Thomas paid them 40 cents a cubic yard for this work. How many cubic yards were taken out in digging the hole?

Test B—SRA. How many cubic yards are taken out in digging a hole 40 feet long, 24 feet wide, and 8 feet deep?

Test C—TIC. A crew of 4 men working for 8 hours with a shovel dug a basement $40 \times 24 \times 8$ for which Mr. Thomas paid them 40 cents a yard. How many yards were excavated?

Test D—SRC. Mr. Thomas hired men to dig a basement 40 feet long, 24 feet wide, and 8 feet deep. How many cubic yards of dirt were taken out in digging the hole?

Problem 9

Test A—TRC. The list price of caps is $3.50. During a sale, Walter bought a cap on which a discount of 15% was given. What is the net price Walter paid?

Test B—TRA. The list price is $3.50; discount is 15%. What is the net price?

Test C—SIC. The Johnson Clothing Company gave a reduction of 15% on all goods bought on June 12. At the sale, Walter bought a cap which usually sold for $3.50; a shirt, for $4.50; and a tie for $1.50. He gave the clerk a ten-dollar bill. What did Walter pay for the cap bought on June 12?

Test D—SRA. One article usually sells for $3.50. It was sold for 15% less than the usual price. What price was paid for the article?

Problem 10

Test A—SRC. Mr. Jones shipped to his agent in Chicago, a carload of peaches, 410 bushels, which he sold at $1.25 a bushel. The agent received 8% of the selling price for selling the peaches. The other expenses were: freight, $75.65; drayage at Chicago, $8.50; baskets, $49.20; picking and packing, $50.20; carting to the station, $18; refrigeration, $55. How much a bushel did Mr. Jones receive after expenses were taken from the amount the agent received for the peaches?

Test B—TRA. A commission of 8% was paid to an agent who sold 410 bu. of fruit at $1.25 per bu. Other expenses were $75.65, $8.50, $49.20, $50.20, $18, and $55. How much per bu. net was received?

Test C—SIA. The sale of 410 bushels at $1.25 a bushel was made by an agent who usually sold a total of $200,000 worth each year. He received as his share 8% of this amount collected. Other expenses were $75.65, $8.50, $49.20, $50.20, $18 and $55. How much did the owner receive for one bushel after all expenses were taken from the amount collected?

Test D—TIA. The sale of 410 bushels at $1.25 a bushel was made by an agent who usually sold a total of $200,000 worth each year. A commission of 8% was paid. Other expenses were $75.65, $8.50, $49.20, $50.20, $18 and $55. How much per bushel net was received?

Problem 11

Test A—TIC. Mr. Thomas deposited $800 in a savings account and $325 in a checking account with the First National Bank, which pays interest at the rate of 4% compounded semiannually. How much will Mr. Thomas have in his savings account at the end of two years?

Test B—SRC. Mr. Thomas deposited $800 in a savings account with The First National Bank, which pays 4 per cent interest every six months. The interest, when due, is added to the amount deposited. How much will he have at the end of two years?

Test C—TRC. Mr. Thomas deposited $800 in a savings account with the First National Bank which pays 4% interest compounded semiannually. How much will he have at the end of two years?

Test D—TRA. A man deposited $800 in a savings account of a bank which pays 4% interest compounded semiannually. How much will be due at the end of two years?

Problem 12

Test A—SRA. A man had an average income last year of $79.17 a month. His average monthly expenses were $60.50. If he had $632.80 at the beginning of the year, how much did he have at the end of the year?

Test B—TIA. A man had an average income last year of $79.17 a month. His average monthly expenses were $60.50 of which $350 was spent each year for food and clothing. If he started the year with a balance on hand of $632.80, what was the balance at the end of the year?

Test C—SIC. Mr. Williams had an average income of $79.17 a month. His average monthly expenses were $60.50 of which $350 was spent for food and clothing. At the beginning of the year Mr. Williams had $632.80 which he had saved previously. How much did he have at the end of the year?

Test D—SIA. A man had an average income of $79.17 a month. His average monthly expenses were $60.50 of which $350 was spent for food and clothing. At the beginning of the year he had $632.80 which he had saved previously. How much did he have at the end of the year?

APPENDIX B

TABLES SHOWING DETAILS OF VARIATIONS IN PUPIL RESPONSES TO DIFFERENT TYPES OF PROBLEM STATEMENTS

The following tables give for each of three grades—sixth, seventh, and eighth—the data relative to the differences in pupil responses to six types of problem statements. For the meanings of the symbols used, see page 20.

Table VI. Variations in Pupil Responses to Problems When Technical Terminology (T) was Substituted for Simplified Terminology (S)—Grade VI

Problems Compared			No. of Pupils		Per Cent Attempted			Per Cent Correct in Principle		
Simple	Technical	Constant Factors	S	T	S	T	Diff.	S	T	Diff.
A 2	B 2	RC	199	185	59.2	59.4	− 0.2	0.0	0.0	0.0
C 5	D 5	RC	198	193	85.3	77.7	7.6	6.5	8.6	− 2.1
B11	C11	RC	185	198	54.0	48.4	5.6	0.0	0.0	0.0
D 4	B 4	CI	193	185	60.1	41.6	18.5	26.7	22.1	4.6
A 8	C 8	CI	199	198	57.7	58.5	− 0.8	15.7	11.2	4.5
A 6	C 6	IA	199	198	85.4	81.8	3.6	17.6	9.3	8.3
C10	D10	IA	198	193	36.8	41.4	− 4.6	4.1	3.8	0.3
D12	B12	IA	193	185	46.1	49.7	− 3.6	6.7	5.4	1.3
C 3	A 3	RA	198	199	97.4	95.4	2.0	20.7	30.0	− 9.3
D 9	B 9	RA	193	185	67.3	56.7	10.6	13.8	15.2	− 1.4

TABLE VII. VARIATIONS IN PUPIL RESPONSES TO PROBLEMS WHEN TECHNICAL TERMINOLOGY (T) WAS SUBSTITUTED FOR SIMPLIFIED TERMINOLOGY (S)—GRADE VII

Problems Compared			No. of Pupils		Per Cent Attempted			Per Cent Correct in Principle		
Simple	Technical	Constant Factors	S	T	S	T	Diff.	S	T	Diff.
A 2	B 2	RC	1485	1487	84.5	84.3	0.2	0.9	0.9	0.0
C 5	D 5	RC	1465	1465	90.5	85.9	4.6	6.3	6.8	− 0.5
B11	C11	RC	1487	1465	47.3	50.6	− 3.3	1.4	3.9	− 2.5
D 4	B 4	CI	1465	1487	82.8	67.1	15.7	55.4	49.4	6.0
A 8	C 8	CI	1485	1465	55.7	61.8	− 6.1	28.3	19.6	8.7
A 6	C 6	IA	1485	1465	79.3	70.8	8.5	32.5	15.5	17.0
C10	D10	IA	1465	1465	44.7	44.3	0.4	6.3	5.9	0.4
D12	B12	IA	1465	1487	34.7	32.5	2.2	14.3	6.8	7.5
C 3	A 3	RA	1465	1485	98.2	98.0	0.2	32.4	40.6	− 8.2
D 9	B 9	RA	1465	1487	72.7	72.2	0.5	50.1	40.9	9.2

TABLE VIII. VARIATIONS IN PUPIL RESPONSES TO PROBLEMS WHEN TECHNICAL TERMINOLOGY (T) WAS SUBSTITUTED FOR SIMPLIFIED TERMINOLOGY (S)—GRADE VIII

Problems Compared			No. of Pupils		Per Cent Attempted			Per Cent Correct in Principle		
Simple	Technical	Constant Factors	S	T	S	T	Diff.	S	T	Diff.
A 2	B 2	RC	644	639	85.4	89.8	− 4.4	5.3	5.4	− 0.1
C 5	D 5	RC	650	647	92.0	87.9	4.1	15.7	18.8	− 3.1
B11	C11	RC	638	650	51.8	54.9	− 3.1	5.1	10.4	− 5.3
D 4	B 4	CI	647	639	86.9	70.6	16.3	73.1	61.2	11.9
A 8	C 8	CI	644	650	63.4	67.8	− 4.4	43.9	32.2	11.7
A 6	C 6	IA	644	650	77.5	63.5	14.0	43.3	30.3	13.0
C10	D10	IA	650	647	53.1	49.9	13.7	14.8	10.2	4.6
D12	B12	IA	647	639	33.8	37.6	− 3.8	26.0	14.8	11.2
C 3	A 3	RA	650	644	99.4	99.1	0.3	43.0	58.5	−15.5
D 9	B 9	RA	647	639	75.0	76.8	− 1.8	65.8	59.1	6.7

TABLE IX. VARIATIONS IN PUPIL RESPONSES TO PROBLEMS WHEN IRRELEVANT DATA (I) WERE SUBSTITUTED FOR RELEVANT DATA (R)—GRADE VI

Problems Compared			No. of Pupils		Per Cent Attempted			Per Cent Correct in Principle		
Relevant	Irrelevant	Constant Factors	R	I	R	I	Diff.	R	I	Diff.
C 5	B 5	SC	198	185	85.4	74.5	10.9	6.5	6.5	0.0
D 8	A 8	SC	193	199	65.2	57.7	7.5	8.7	15.7	− 7.0
C 4	A 4	TA	198	199	44.9	45.7	− 0.8	16.9	15.4	1.5
B10	D10	TA	185	193	40.0	41.4	− 1.4	4.1	3.8	0.3
C 3	B 3	SA	198	185	97.4	88.1	9.3	20.7	23.3	− 2.6
A12	D12	SA	199	193	59.2	46.1	13.1	9.3	6.7	2.6
B 6	D 6	CT	185	193	78.3	79.8	− 1.5	73.8	5.8	68.0
C11	A11	CT	198	199	48.4	42.2	6.2	0.0	0.0	0.0

TABLE X. VARIATIONS IN PUPIL RESPONSES TO PROBLEMS WHEN IRRELEVANT DATA (I) WERE SUBSTITUTED FOR RELEVANT DATA (R)—GRADE VII

Problems Compared			No. of Pupils		Per Cent Attempted			Per Cent Correct in Principle		
Relevant	Irrelevant	Constant Factors	R	I	R	I	Diff.	R	I	Diff.
C 5	B 5	SC	1465	1487	90.5	87.6	2.9	6.3	4.0	2.3
D 8	A 8	SC	1465	1485	65.8	55.7	10.1	25.5	28.3	− 2.8
C 4	A 4	TA	1465	1485	63.0	62.5	0.5	38.1	49.4	−11.3
B10	D10	TA	1487	1465	50.1	44.3	5.8	5.1	5.9	− 0.8
C 3	B 3	SA	1465	1487	98.2	90.7	7.5	32.4	29.2	3.2
A12	D12	SA	1485	1465	35.5	34.7	0.8	19.5	14.3	5.2
B 6	D 6	CT	1487	1465	72.5	68.1	4.4	68.6	9.0	59.6
C11	A11	CT	1465	1485	50.6	36.8	13.8	3.9	1.6	2.3

TABLE XI. VARIATIONS IN PUPIL RESPONSES TO PROBLEMS WHEN IRRELEVANT DATA (I) WERE SUBSTITUTED FOR RELEVANT DATA (R)—GRADE VIII

Problems Compared			No. of Pupils		Per Cent Attempted			Per Cent Correct in Principle		
Relevant	Irrelevant	Constant Factors	R	I	R	I	Diff.	R	I	Diff.
C 5	B 5	SC	650	639	92.0	90.6	1.4	15.7	10.5	5.2
D 8	A 8	SC	647	644	72.0	63.4	8.6	52.4	43.9	8.5
C 4	A 4	TA	650	644	68.5	66.8	1.7	47.2	62.3	−15.1
B10	D10	TA	639	647	55.8	49.9	5.9	13.8	10.2	3.6
C 3	B 3	SA	650	639	99.4	92.0	7.4	43.0	42.9	0.1
A12	D12	SA	644	647	36.5	33.8	2.7	33.6	26.0	7.6
B 6	D 6	CT	639	647	78.1	67.1	11.0	43.3	16.4	26.9
C11	A11	CT	650	644	54.9	38.0	16.9	10.4	7.8	2.6

TABLE XII. VARIATIONS IN PUPIL RESPONSES WHEN LANGUAGE WHICH IS ABSTRACT (A) IN NATURE WAS SUBSTITUTED FOR LANGUAGE WHICH PRESENTS A REAL OR CONCRETE (C) SETTING—GRADE VI

Problems Compared			No. of Pupils		Per Cent Attempted			Per Cent Correct in Principle		
Concrete	Abstract	Constant Factors	C	A	C	A	Diff.	C	A	Diff.
A 2	C 2	SR	199	198	59.2	57.1	2.1	0.0	0.9	− 0.9
D 8	B 8	SR	193	185	65.2	56.7	8.5	8.7	13.3	− 4.6
B 4	A 4	TI	185	199	41.6	45.7	− 4.1	22.1	15.4	6.7
D 6	C 6	TI	193	198	79.7	81.8	− 2.1	5.8	9.3	− 3.5
B 5	A 5	SI	185	199	74.5	73.8	0.7	6.5	6.8	− 0.3
C12	D12	SI	198	193	53.0	46.1	6.9	7.6	6.7	0.9
A 9	B 9	RT	199	185	60.3	56.8	3.5	20.0	15.2	4.8
C11	D11	RT	198	193	48.5	48.7	− 0.2	0.0	0.0	0.0

TABLE XIII. VARIATIONS IN PUPIL RESPONSES WHEN LANGUAGE WHICH IS ABSTRACT (A) IN NATURE WAS SUBSTITUTED FOR LANGUAGE WHICH PRESENTS A REAL OR CONCRETE (C) SETTING—GRADE VII

Problems Compared			No. of Pupils		Per Cent Attempted			Per Cent Correct in Principle		
Concrete	Abstract	Constant Factors	C	A	C	A	Diff.	C	A	Diff.
A 2	C 2	SR	1485	1465	84.5	75.4	9.1	0.9	0.8	0.1
D 8	B 8	SR	1465	1487	65.8	58.8	7.0	25.5	25.6	− 0.1
B 4	A 4	TI	1487	1485	67.1	62.5	4.6	49.4	49.4	0.0
D 6	C 6	TI	1465	1465	68.1	70.8	− 2.7	9.0	15.5	− 6.5
B 5	A 5	SI	1487	1485	87.6	90.8	− 3.2	4.0	3.3	0.7
C12	D12	SI	1465	1465	35.2	34.7	0.5	15.0	14.3	0.7
A 9	B 9	RT	1485	1487	72.3	72.2	0.1	50.4	40.9	9.5
C11	D11	RT	1465	1465	50.6	50.6	0.0	3.9	2.8	1.1

TABLE XIV. VARIATIONS IN PUPIL RESPONSES WHEN LANGUAGE WHICH IS ABSTRACT (A) IN NATURE WAS SUBSTITUTED FOR LANGUAGE WHICH PRESENTS A REAL OR CONCRETE (C) SETTING—GRADE VIII

Problems Compared			No. of Pupils		Per Cent Attempted			Per Cent Correct in Principle		
Concrete	Abstract	Constant Factors	C	A	C	A	Diff.	C	A	Diff.
A 2	C 2	SR	644	650	85.4	81.4	4.0	5.3	6.2	− 0.9
D 8	B 8	SR	647	639	72.0	63.7	8.3	52.4	49.9	2.5
B 4	A 4	TI	639	644	70.6	66.8	3.8	61.2	62.3	− 1.1
D 6	C 6	TI	647	650	67.1	63.5	3.6	16.4	30.3	−13.9
B 5	A 5	SI	639	644	90.6	91.5	− 0.9	10.5	9.2	1.3
C12	D12	SI	650	647	36.6	33.8	2.8	23.9	26.0	− 2.1
A 9	B 9	RT	644	639	72.2	76.8	− 4.6	66.9	59.1	7.8
C11	D11	RT	650	647	54.9	52.4	2.5	10.4	13.6	− 3.2